Distribution Cumulée (le cas discret)

Définition (les effectifs cumulés croissants (Fréquences

- L'effectif cumulé croissant noté $N_i^{\nearrow}$ est la somme des effectifs correspondants aux valeurs du caractère inferieur ou égale à x_i

d'où $N_i^{\nearrow} = \sum_{p=1}^{i} n_p$

$F_i^{\nearrow}$: La fréquence cumulée croissante Tq :

$$F_i^{\nearrow} = \sum_{p=1}^{i} f_p .$$

Définition : (les effectifs cumulées décroissants)

- L'effectif cumulé décroissant $N_i^{\searrow}$ est la somme des effectifs correspondants aux valeurs du caractère superieures ou égale à x_i

d'où $N_i^{\searrow} = \sum_{p=i}^{k} n_p \quad 1 \leq i \leq k$

$F_i^{\searrow}$: la fréquence cumulée decroissante Tq :

$$F_i^{\searrow} = \sum_{p=i}^{k} f_p .$$

Exemple

x_i	n_i	f_i	$N_i\nearrow$	$F_i\nearrow$	$N_i\searrow$	$F_i\searrow$
0	16	0,25	16	0,25	64 = n	1
1	18	0,291	34	0,541	48	0,75
2	14	0,218	48	0,759	30	0,459
3	11	0,172	59	0,931	16	0,241
4	3	0,047	62	0,978	5	0,069
5	2	0,031	64 = n	1	2	0,022
$\sum$	64					

$$N_i\nearrow = \sum_{p=1}^{i} n_p$$

$N_1 = n_1 = 16$

$N_2 = \sum_{p=1}^{2}(n_1+n_2)$

$= 16+18 = 34$

$$N_i\searrow = \sum_{p=1}^{k=6} n_p$$

$N_1\searrow = \sum_{p=1}^{6}(n_1+n_2+n_3+n_4+n_5+n_6)$

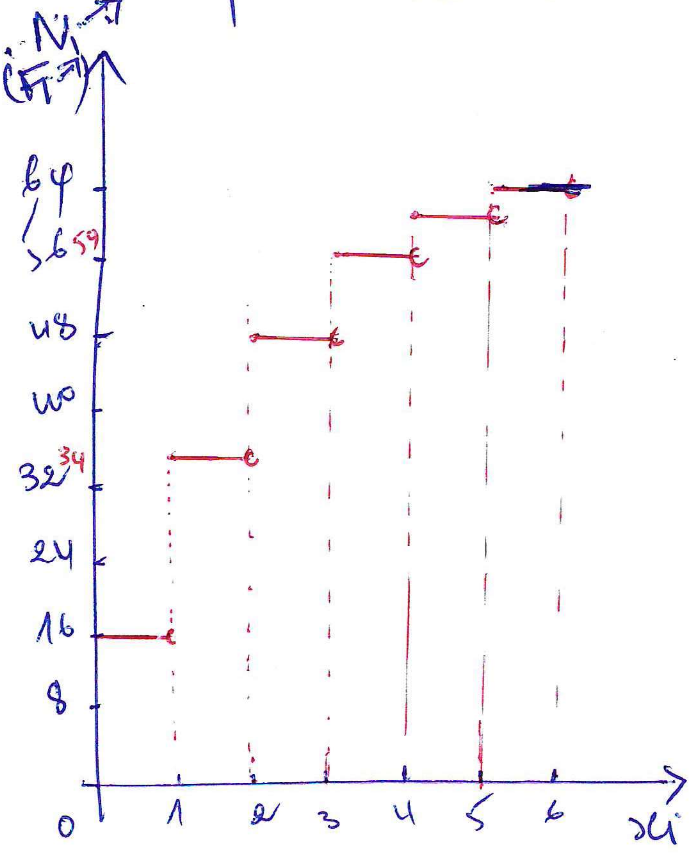
$= 64$

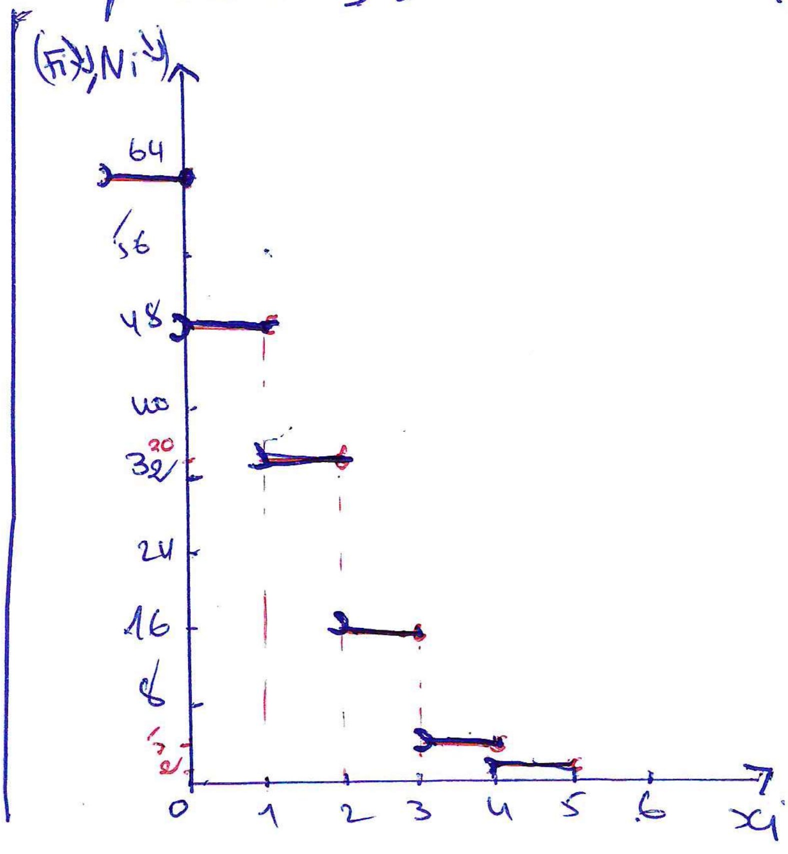
$N_2\searrow = \sum_{2}^{6}(n_2+n_3+n_4+n_5+n_6)$

$= 48$

La même chose pour $F_i\searrow$

- **Représentation graphique des** $(N_i\nearrow, N_i\searrow, F_i\nearrow, F_i\searrow)$ c'est un graphe en escaliers dont les marches correspondants au valeurs possibles du caractère x_i



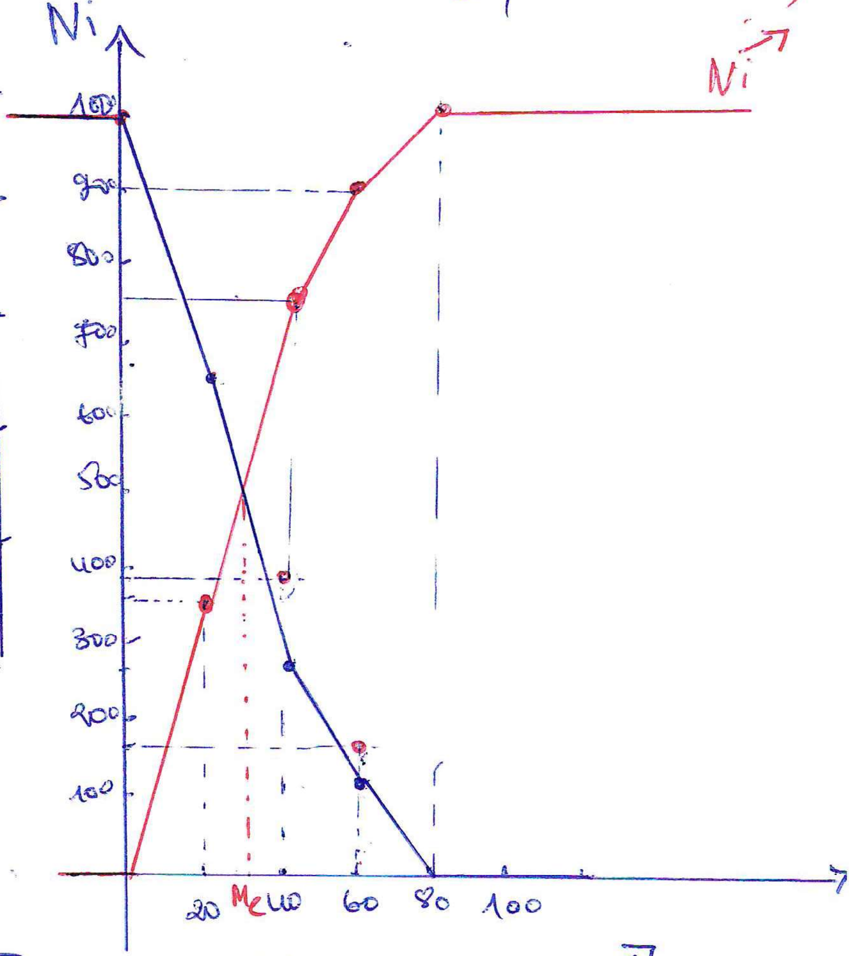


Remarque : Dans le cas continu, on définit d'une manière analogue les effectifs ou fréquences cumulés croissant en prenant à la place de x_i la borne supérieur de classe $[e_i, e_{i+1}[$

- Pour les effectifs ou fréquences cumulés décroissants en prenant à la place de x_i la borne inférieur de la ~~to~~ classe $[e_i, e_{i+1}[$.

Exemple : Représentation graphique ~~des~~ effectifs cumulés croissants et décroissants (cas continu)

classe	n_i	$N_i^{\nearrow}$	$N_i^{\searrow}$	f_i
$[0, 20[$	360	360	1000	0,36
$[20, 40[$	380	740	640	0,38
$[40, 60[$	160	900	260	0,16
$[60, 80[$	100	1000	100	0,1



courbe cumulées des effectifs croissants et décroissants

$[0, \underline{\underline{20}}[$ $N_i^{\nearrow}$

$[\underline{\underline{0}}, 20[$ $N_i^{\searrow}$

Paramètres caractéristiques : pour l'étude d'un caractère quantitatif on s'intéresse à deux types de paramètres :

a) Les caractéristiques de tendance centrale ou de position : sont des valeurs qu'on retrouve au centre des observations.

b) Les caractéristiques de dispersion : indiquent la disposition des observations autour de la moyenne.

a) Les caractéristique de Position :

1) Le mode :

1.1 cas discret : c'est la valeur du caractère quantitatif discret qui a l'effectif ou fréquence le plus élevée et on le note par M_o.

Exemple :

x_i	0	1	2	3	4	5
n_i	16	18	14	11	3	2

. Le mode : $M_o = 1$, $n_{i\,max} = 18$.

1.2 Cas continu : dans ce cas, on estime M_o par la classe modale "ayant l'effectif le plus élevé.

Exemple :

classe	[0,20[	[20,40[	[40,60[	[60,80[
n_i	360	380	160	100

La classe modale st [20,40[car $n_{i_{max}} = 380$

- On peut appeler M_o (mode) le centre de classe modale i.e $M_o = \frac{20+40}{2} = 30$.
- Pour du calcul approché, on utilise la formule suivante : Soit la classe modale $[e_{i-1} - e_i[$

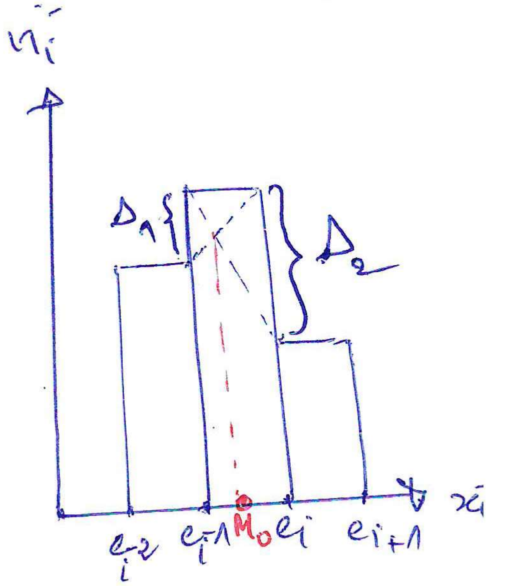
On a $M_o \in [e_{i-1}, e_i[$.

$$M_o = e_{i-1} + a_i \frac{\Delta_1}{\Delta_1 + \Delta_2}$$ avec :

a_i : l'amplitude de la classe $[e_{i-1} - e_i[$

Δ_1 : l'excès de la classe modale par rapport à la classe précédent

Δ_2 : " " " suivante.



Remarque : on peut avoir un seul mode ou plusieurs modes alors la série est multimode.

exemple : $M_o = 20 + 20 . \frac{380 - 360}{20 + 220}$

$M_o = 21,66$

2) La Médiane :

2.1 Cas discret : Considérons la série quantitative discrète dont les valeurs sont ordonnées dans le sens croissant $x_1, x_2, \ldots x_n$.

La Médiane notée M_e est la valeur qui divise la série originale en deux séries partielles égales.

On distingue deux cas :

a) si n impair $\quad M_e = x\left(\frac{n+1}{2}\right)$

b) si n pair $\quad M_e = \dfrac{x(\frac{n}{2}) + x(\frac{n}{2}+1)}{2}$

exemple : soit la série suivante :

2, 2, 3, 3, 3, 4, 5, 6, 6, 7, 8, 8, 9 ; $n = 13$.

n est impair $\Rightarrow M_e = x\left(\frac{n+1}{2}\right) = x_7 = 5$

$\boxed{M_e = 5}$

. Soit la série suivante :

2, 2, 3, 3, 3, 4, $\underset{x_7}{5}$, $\underset{x_8}{6}$, 6, 7, 8, 8, 9, 9 $\quad n = 14$.

n est pair $\Rightarrow M_e = \dfrac{x(n/2) + x(n/2+1)}{2} = \dfrac{x_7 + x_8}{2} = 5{,}5$

exemple :

$\underbrace{0, 0 \ldots 0}_{16}, \underbrace{1, 1 \ldots 1}_{18}, \underbrace{2, 2, \ldots 2}_{14}, \underbrace{3, 3 \ldots 3}_{11}, \underbrace{4, 4 \ldots 4}_{13}, \underbrace{5, 5 \ldots 5}_{2}$ $\quad \boxed{n = 64}$

$n = 64$ pair $\Rightarrow M_e = \dfrac{x(\frac{n}{2}) + x(\frac{n}{2}+1)}{2} = \dfrac{x_{32} + x_{33}}{2}$

$= \dfrac{1+1}{2} = 1$

$\boxed{M_e = 1}$

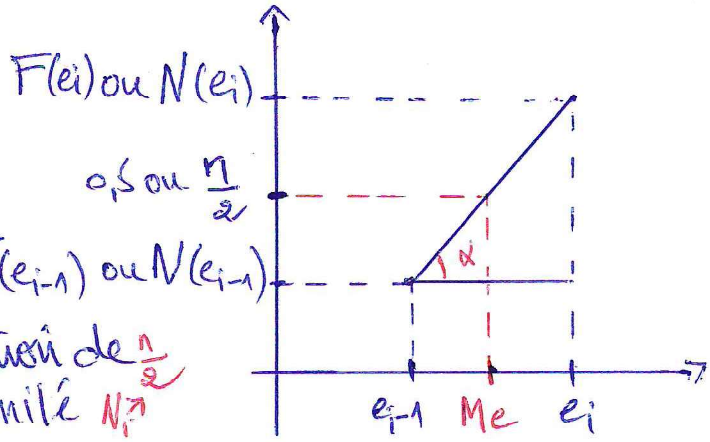
Remarque :

Peut être $M_e \notin$ aux valeurs de la série (n'appartient pas)

- On peut utiliser une autre méthode pour déterminer M_e, On calcul les $N_i \nearrow$ ou $F_i \nearrow$ et on cherche les deux valeurs qui encadrent $\frac{n}{2}$ ou 0,5, x_i qui correspond à la plus grande valeur d'entre eux nous donne Me.

2.2 Cas Continu:

- Géométriquement, la médiane est la projection du point de l'interssection de deux courbes Cumulatives croissante et décroissante des effectifs (ou fréquences) sur l'axe (ox).
- Pour déterminer la Médiane M_e, on utilise la Méthode d'interpolation linéaire :



- On determine la classe qui contient M_e, on cherche la position de $\frac{n}{2}$ par rapport aux effectifs cumilé $N_i \nearrow$ et la plus grand valeur.

$$tg\alpha = \frac{\frac{n}{2} - N(e_{i-1})}{M_e - e_{i-1}} = \frac{N(e_i) - N(e_{i-1})}{e_i - e_{i-1}}$$

$$M_e = e_{i-1} + \frac{e_i - e_{i+1}}{N(e_i) - N(e_{i+1})}\left(\frac{n}{2} - N(e_{i-1})\right)$$

$$M_e = e_{i-1} + a_i \frac{\frac{n}{2} - N(e_{i-1})}{n_i}$$

a_i: amplitude
n_i: effectif de la classe médiane

Exemple :

classe	n_i	$N_i \nearrow$	f_i	$F_i \nearrow$
[0,20[	360	360	0,36	
[20,40[	380	740	0,38	
[40,60[	160	900	0,16	
[60,80[	100	1000	0,1	

$n = 1000$

$\frac{n}{2} = 500$

$360 < 500 < 740$

$740 \to M_e \in [20,40[$

$e_{i-1} = 20$

$e_i = 40$

$N(e_i) = 740$

$N(e_{i-1}) = 360$

D'après la formule

$$M_e = e_{i-1} + a_i \frac{\frac{n}{2} - N(e_{i-1})}{N(e_i) - N(e_{i-1})}$$

$$= 20 + 20 \cdot \frac{(500 - 360)}{740 - 360} \simeq 27,37$$

$M_e = 27,37 \in [20,40[$

$$M_e = e_{i-1} + a_i \frac{0,5 - F(e_{i-1})}{F(e_i) - F(e_{i-1})}$$

3. La Moyenne :

3.1 La Moyenne arithmétique simple :

$$\text{Moyenne arithmétique simple} = \frac{\text{Somme des donnés}}{\text{Nombre de données}}$$

Exemple : si les notes d'un étudiant sont de : 10, 15, 8, alors la moyenne arithmétique simple est de : $\frac{10+15+8}{3} = 11$.

3.2 La Moyenne arithmétique pondérée :

Si les notes au Baccalauréat d'un élève sont de 11/20 en français (coefficient 5), de 13/20 en histoire (coefficient 3) et de 10/20 en Maths (coefficient 7), sa moyenne arithmétique pondérées est de :

$$\bar{X} = \frac{(11\times5)+(13\times3)+(10\times7)}{(5+3+7)} = 10,93.$$

donc $\bar{X} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$ ou $\bar{X} = \sum_{i=1}^{k} f_i x_i$

Rq : Dans le cas continu en prenant les x_i comme centre des classes.

Exemple :

x_i	n_i	$n_i x_i$
0	16	0
1	18	18
2	14	28
3	11	33
4	3	12
5	2	10
Σ	64	101

$$\bar{X} = \frac{1}{n}\sum n_i x_i = \frac{101}{64} = 1,58$$

3.3 La Moyenne quadratique :

La Moyenne quadratique MQ d'une série de n valeurs $x_1, x_2 \ldots x_n$, corres pondent respectivement les effectifs $n_1, n_2 \ldots n_k$ est égale :

$$MQ = \sqrt{\frac{1}{n}\sum_{i=1}^{k} n_i x_i^2}$$

Exemple :

La moyenne quadratique MQ des valeurs : $-2, 5, -8, 9, -4$ est

$$MQ = \sqrt{\frac{1}{5}\left((-2)^2+(5)^2+(-8)^2+(9)^2+(-4)^2\right)} = 6,16$$

3.4 La moyenne géométrique : est la racine nième défini comme suit :

$$\bar{G} = \sqrt[n]{x_1 \times x_2 \times x_3 \ldots x_n}$$

3.5 La moyenne harmonique :

$$H = \frac{1}{\frac{1}{n}\sum_{1}^{k}\frac{n_i}{x_i}} = \frac{n}{\sum_{i=1}^{k}\frac{n_i}{x_i}}$$

Exemple : H des 1, 4, 8, 10, 12 est

$$H = \frac{5}{\frac{1}{1}+\frac{1}{4}+\frac{1}{8}+\frac{1}{10}+\frac{1}{12}} = 3,2$$

Remarque : On a la relation suivante entre ces différentes moyennes.

$$H \leq G \leq \bar{x} \leq MQ$$

4. Les Quartiles : sont les valeurs qui divisent une série discrète ordonnée dans le sens croissant en 4 partie égales avec :

Q_1 : ~~est~~ la valeur dont l'indice est le plus petit entier supérieur à $\left(\frac{n}{4}\right)$

$Q_2 = Me$

Q_3 : ~~est~~ la valeur dont l'indice est le plus petit entier supérieur à $\left(\frac{3n}{4}\right)$

- L'écart interquartile est donné par : $Q_3 - Q_1$

Exemple : $0 \underset{16}{—} 0 \quad 1 \underset{18}{—} 1 \quad 2 \underset{14}{—} 2 \quad 3 \underset{11}{—} 3 \quad 4 \underset{3}{—} 4 \quad 5 \underset{2}{—} 5$ $\quad n = 64$

$Q_2 = Me$, $Q_1 = \cancel{x\left(\frac{n}{4}\right)} = ?!$ $\frac{n}{4} = 16 \Rightarrow Q_1 = x_{(17)} = 1$

$Q_3 = ?!$ $\frac{3n}{4} = 48 \Rightarrow Q_3 = x(49) = 3$

- **Les déciles** : qui divisent la série discrète ordonnée dans le sens croissant en 10 partie :

D_1 : est la valeur dont l'indice est le plus petit entier supérieur à $\frac{n}{10}$

D_2 : ———— $\frac{2n}{10}$

⋮

D_9 : ———— $\frac{9n}{10}$

- **Les centiles** :

C_1 ———— $\frac{n}{100}$

C_2 ———— $\frac{2n}{100}$

⋮

C_{99} ———— $\frac{99n}{100}$

Détermination des quartiles:

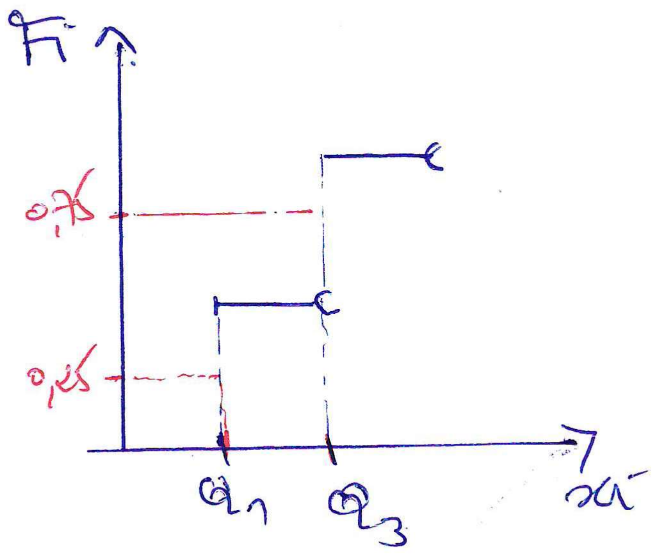
Cas continu: Pour déterminer les quartiles, on utilise la même méthode d'interpolation linéaire pour la médiane: (on détermine la classe de Q_1, Q_2, Q_3)

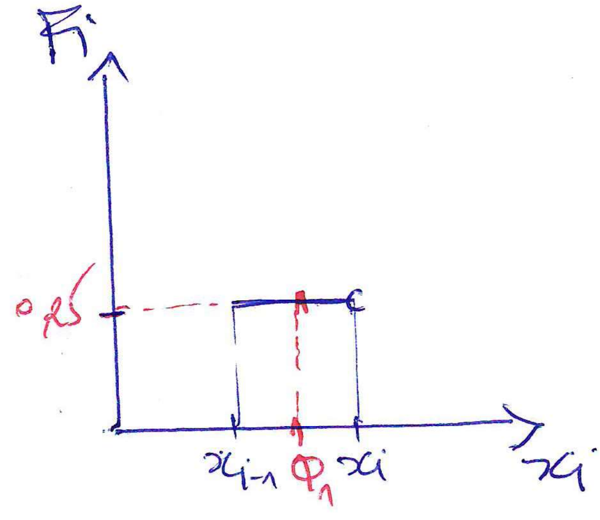
décile

- $Q_1 = e_{i-1} + (e_i - e_{i-1}) \dfrac{\frac{n}{4} - N(e_{i-1})}{N(e_i) - N(e_{i-1})}$
- $Q_2 = Me = e_{i-1} + (e_i - e_{i-1}) \dfrac{\frac{n}{2} - N(e_{i-1})}{N(e_i) - N(e_{i-1})}$
- $Q_3 = e_{i-1} + (e_i - e_{i-1}) \dfrac{\frac{3n}{4} - N(e_{i-1})}{N(e_i) - N(e_{i-1})}$

- Géométriquement Q_1, Q_3 est la projection du point $\frac{n}{4}, \frac{3n}{4}$ respectivement de courbe cumulative croissante sur l'axe (ox).

Cas discret: (graphiquement)





$$Q_1 = \frac{x_i + x_{i-1}}{2}$$